Utility of an Observational Social Skill Assessment as a Measure of Social Cognition in Autism

Grace Lee Simmons et al.

Citation of published version:
Simmons, G. L., Ioannou, S., Smith, J. V., Corbett, B. A., Lerner, M. D., & White, S. W. (2020). Utility of an Observational Social Skill Assessment as a Measure of Social Cognition in Autism. In Autism Research (Vol. 14, Issue 4, pp. 709–719). Wiley. https://doi.org/10.1002/aur.2404

HHS Public Access

Author manuscript

Autism Res. Author manuscript; available in PMC 2022 April 01.

Published in final edited form as:

Autism Res. 2021 April ; 14(4): 709–719. doi:10.1002/aur.2404.

Utility of an Observational Social Skill Assessment as a Measure of Social Cognition in Autism

Grace Lee Simmons, M.A. [Graduate Clinician and Researcher],
Center for Youth Development and Intervention, University of Alabama, Department of Psychology, Box 870348, Tuscaloosa, AL 35487

Sara Ioannou, M.A. [Research Assistant],
Department of Psychiatry & Behavioral Sciences, Vanderbilt University Medical Center, Nashville, TN, 37212

Jessica V. Smith [Student],
Department of Psychology, University of Alabama, Box 870348, Tuscaloosa, AL 35487

Blythe A. Corbett, Ph.D. [Professor of Psychiatry & Behavioral Sciences],
Professor of Psychology, Vanderbilt University Medical Center, Nashville, TN, 37212

Matthew D. Lerner, Ph.D. [Associate Professor, Clinical Psychology],
Department of Psychology, Stony Brook University, Stony Brook, NY, 11794

Susan W. White, Ph.D., ABPP [Doddridge Saxon Chairholder in Clinical Psychology]
Center for Youth Development and Intervention, Department of Psychology, University of Alabama, Box 870348, Tuscaloosa, AL 35487

Abstract

Models of impaired social competence in Autism Spectrum Disorder (ASD) highlight deficits in social cognition and social behavior. The Contextual Assessment of Social Skills (CASS) is a laboratory-based assessment of conversation ability in which participants interact with trained confederates who act interested (CASS-I) and bored (CASS-B), sequentially. The increased ecological validity of the CASS allows for better generalization to real-world social situations. Participants' perceptions of confederate behavior, assessed by the CASS Conversation Rating Scale (CRS), might offer additional utility as a metric of social cognition. The current study examined CASS confederate behavior (adherence to interested or bored condition) and both internal validity and convergent validity of the CASS as a measure of social behavior and social cognition. Fifty adolescents with ASD participated as part of a multi-site randomized clinical trial. Adherence ratings were consistent across gender and site, with interested confederates significantly out-performing bored confederates. The ability to distinguish between interested and bored confederates was positively associated with CASS social behavior and social cognition tasks, although social behavior during the CASS was not consistently associated with parent-rated social behavior. Controlling for confederate behavior did not significantly alter these associations. Findings demonstrate strong internal validity of the CASS and, partially, external validity of the CASS as a measure of social cognition. Findings highlight nuanced differences in social behavior

Corresponding Author: Grace Lee Simmons, M.A., (205)-348-4201, GLSimmons@crimson.ua.edu.

and social cognition. The CASS shows promise as an outcome measure for clinical interventions and should be incorporated into a multi-method battery to assess social competence in individuals with ASD.

Lay Abstract

Social cognition and social behavior should be studied together to examine social competence in youth with autism. The CASS, a behavioral observation measure, shows promise towards this end; findings suggest the CASS taps social cognition and social behavior when administered alongside a participant rating scale of their conversation partner's engagement. Continued research, including examination of the CASS, may inform best practices in comprehensive assessment of social competence in Autism.

Keywords

Adolescents; Children; Social Cognition; Social Cognition & Theory of Mind; Face Perception

Social competence, understood as the ability to navigate the social world and maintain healthy, functional relationships across domains (e.g., home, school, employment, intimacy, etc., see Lemerise & Arsenio, 2000), is recognized as an area of core diagnostic impairment in individuals with Autism Spectrum Disorder (ASD; American Psychiatric Association, 2013). There is not, however, consensus regarding an etiological model of impaired social competence in ASD. Research has focused on the social brain (e.g., Kennedy & Adolphs, 2012); social amotivation (Chevallier et al., 2012); impairments in social cognition (i.e., impaired cognitive processes underlying social behavior, particularly difficulty recognizing faces and facial expressions; Velikonja et al., 2019); and atypical social behavior (e.g., poor eye contact, reduced use of gestures, difficulty with reciprocal conversations; American Psychiatric Association, 2013). Underlying these approaches is the common understanding that social cognition is indelibly linked to social behavior, and while these models suggest an association between these constructs, they do not support direct correspondence (Pallathra et al., 2018). As such, social cognition and social behavior must be studied in tandem to assess and intervene on social competence in individuals with ASD (e.g., Corbett et al., 2016, see Figure 1).

Numerous standardized instruments assess components of social cognition such as facial emotion recognition, memory for faces, theory of mind, etc. (see McMahon, Lerner, & Britton, 2013 for review of standardized instruments). However, measurement of social behavior typically relies on parent-report questionnaires (e.g., Social Responsiveness Scale-2, Constantino & Gruber, 2012; Vineland Adaptive Behavior Scale, Sparrow, Cicchetti, & Salunier, 2016, etc.) which are influenced by context and ostensibly reflect the caregiver's perceptions and bias. As a result, there is a growing emphasis on the importance of ecologically valid assessments of social behavior in individuals with ASD (Cordier et al., 2015; Wolstencroft et al., 2018). Naturalistic observations such as school-based and playground observations (e.g., Corbett et al., 2016), while ecologically valid, are resource-intensive and are difficult to integrate into standard assessment batteries and intervention protocols. The Autism Diagnostic Interview Schedule (ADOS-2, Lord et al.,

2012), administered by a rigorously-trained clinician to evaluate reciprocal social behavior in youth with ASD, is also resource-intensive and findings may not translate to social interactions with same-age peers. Instead, laboratory-based observations of social skill are an accessible and more ecologically valid alternative to assess impaired social ability in youth with ASD.

The Contextual Assessment of Social Skills (CASS; Ratto, Turner-Brown, Rupp, Mesibov, & Penn, 2011) is a laboratory-based observational measure of conversation ability which may generalize more directly to real world social skills. The CASS has been used to evaluate social impairment in youth (Dolan et al., 2016) and adults (White et al., 2015) with ASD. The CASS consists of two videotaped conversations between the participant and two trained confederates[1]: the first confederate shows interest (CASS-I) in the conversation and the second confederate behaves as though bored (CASS-B). Participants' behavior in the videos is coded for quantitative data (e.g., number of questions asked) and qualitative data (e.g., positive affect) capturing engagement with the confederate. Coded differences between CASS-I and CASS-B are thought to capture participants' awareness of their conversation partner (i.e., social cognition) and accompanying social behavior. Although the setting (lab-based) is not naturalistic, the CASS is thought to be fairly ecologically valid because the conversational partner is of a similar age and it is structured to pull for casual conversation. Individuals with ASD have demonstrated less engagement with an interested conversation partner compared with neurotypical peers, as reflected in fewer questions asked, fewer topic changes, and lower ratings of rapport and overall involvement in the conversation (Ratto et al., 2011). Additionally, the Conversation Rating Scale (CRS, Ratto et al., 2011), a self-report questionnaire participants complete after each conversation, was developed as a manipulation check to evaluate participants' recognition of the two conditions. Preliminary findings showed individuals with ASD rated bored confederates as less engaged than interested confederates, but the between-condition difference was smaller than was seen in individuals without ASD (Ratto et al., 2011). This might reflect a deficit in social cognition, where autistic youth were less perceptive of bored confederates' social behaviors. This measured change in CRS ratings between interested and bored confederates (CRS-C) could therefore serve as an indicator of underlying social cognition accompanying this behavioral change, though scant research has yet to use the CRS in this way (e.g., Leonczyk, 2017). Additional research focusing on the CRS is needed to evaluate the utility of the CASS as a measure of social cognition.

To the authors' knowledge, research has yet to examine the particular value of CASS-B and CRS in contributing to the CASS's utility as a measure of social behavior and social cognition. The contrast between CASS-I and CASS-B might be understood to reflect a downstream effect of a participant's own social cognition, wherein he/she first interprets the confederate's behavior (i.e., recognition of facial expression, tone of voice, etc.) and responds with appropriate social behaviors (e.g., increased effort to engage the bored confederate; Ratto et al., 2011), demonstrating how social cognition and behavior are independent, yet necessarily linked to each other. In other words, the ability to gauge another

[1]'Confederate' is an established psychological term which refers to actors trained to interact with research study participants (i.e., research helpers; Allen, 2017)

person's social engagement (e.g., theory of mind) ought to drive social behavior. While the initial study showed individuals with and without ASD made more of an effort to engage with a bored confederate compared to an interested confederate across the four targeted domains, those with ASD adjusted their behavior less than did their neurotypical peers (Ratto et al., 2011). Inclusion of both conversational conditions and concurrent examination of the CRS has the potential to provide a valuable, observer-rated behavioral index and accompanying self-report measure of social cognition which is untapped by established measures of parent-reported social behavior and clinician-assessed social cognition.

Furthermore, implementation of the CASS, particularly the use of trained confederates, gives rise to concern about inconsistencies across assessments which may influence participant behavior. Dolan and colleagues (2016) examined confederate behavior during CASS-I, finding two female confederates differed significantly in their rapport with participants; the authors subsequently excluded CASS-I assessments with female confederates from research analyses. This approach, in addition to excluding informative data, fails to consider the effect of differential behavior on the CASS measure itself. Research has yet to examine the robustness of the CASS in the face of confederates who fail to follow the instructions of their respective condition and unduly influence participant behavior; findings would speak to identified implementation concerns regarding the CASS.

The current study had two main aims: to examine CASS confederate behavior as it aligns with administration instructions and to examine internal and convergent validity of the CASS as a measure of social behavior and social cognition in a sample of adolescents with ASD. We hypothesized 1) trained confederates would consistently (e.g., at least 80%) adhere to the prescribed behaviors in their respective conversational conditions. In examining the internal validity of the CASS, and given prior findings of autistic individuals' behavioral change between CASS-I and CASS-B (Ratto et al., 2011; White et al., 2015), we hypothesized 2) behavior during CASS-I would differ significantly from behavior during CASS-B, and 3) behavioral change would be accompanied by a significant difference in participants' insight into confederates' behavior (i.e. CRS ratings). Considering social cognition deficits characteristic of ASD, we also hypothesized 4) autistic participants' ratings of confederate behavior on the CRS would differ significantly from those of a trained observer. Finally, we hypothesized 5) CASS and CRS would demonstrate convergent validity with extant measures of social behavior (i.e., SRS-2 Social Communication Index and ADOS-2 Social Affect Subscale) and social cognition (i.e., NEPSY scales of Affect Recognition, Memory for Faces, Memory for Faces Delayed, and Theory of Mind).

Methods

Participants in this study were enrolled in a multi-site randomized control trial examining the efficacy of a theatre-based intervention in improving social competence in youth with ASD. Participants were recruited via community events, local mental healthcare providers, social media, and contact registries. Participants (ages 10–16) had a diagnosis of ASD (confirmed via ADOS-2), and Full-Scale IQ ≥ 70[2]. All participants and their caregivers consented to participate in the study.

Participants

Sixty-four adolescents initially enrolled, with fourteen excluded from data analyses: three were ineligible; seven withdrew, and four video files were corrupted, leaving a final sample of 50 adolescents (Table 1). Twenty-eight trained confederates (sixteen female, twelve male) ranging in age from 9–20 years old (*M:* 15.50, *SD:* 3.71) participated in CASS assessments across three sites. Confederates were recruited from the community (e.g., flyers, word of mouth) and consented (caregiver consent acquired for those under 18) to participate in this study. Confederates were paid $20 per hour for participation.

Measures

Contextual Assessment of Social Skills (Ratto et al., 2011)—The CASS is a laboratory-based observation measure wherein individuals with ASD participate in two three-minute conversations with unfamiliar confederates of the opposite gender, per the original protocol (Ratto et al., 2011). All confederates participated in a one-hour didactic training which included behavioral instructions and structured rehearsal; confederates were not provided with explicit conversation topics, therefore discussions varied significantly across participants. Additionally, feedback regarding performance during CASS-I and CASS-B was provided during training and after each CASS administration, as needed. Videotapes of conversations were coded on nine criteria assessing conversational skill: questions asked, topic changes, vocal expressiveness, gestures, positive affect, kinesic arousal, social anxiety, overall involvement in the conversation, and overall quality of rapport. Each item was rated on a 7-point Likert scale integrating specific behavioral observations; greater values reflect increased conversational ability/effort, per the original CASS Scoring Manual (Ratto et al., 2011). Coders also participated in structured training, which involved coding eight training videos provided by Ratto and colleagues (2011). Completion of the training and attainment of reliability required achieving 80% agreement overall with ‘Gold Standard’ codes provided by the CASS authors (Ratto et al., 2011) prior to coding CASS data. Each site maintained coding calibration via double-coding of 10% of all videos by trained research personnel. A team of established coders (e.g., graduate student, licensed psychologist) assessed coder drift by double-coding 20% of all study videos across all sites. Intraclass correlations were computed for each of the nine items of the CASS (ICC: .75–.98), reflecting strong cross-site agreement.

To assess confederate adherence to the instructions of their respective conversations (Interested and Bored, operationalized in Ratto et al., 2011), the research team designed a 4-point Likert scale (1 “fully disengaged” to 4 “fully engaged”; Appendix A) which considered both non-verbal gestures (e.g., posture/hand gestures), and verbal cues, (e.g., duration between topic changes). The confederate adherence code was split to create a binary rating of “hit” or “miss” for each CASS conversation: a “hit” indicated the confederate adhered to their prescribed condition; the CASS-I confederate was rated as either 3 or 4 (i.e., more interested than bored), and the CASS-B confederate was rated 1 or 2 (i.e., more bored than interested). A single undergraduate coder was trained first to reliability

[2]Participants with FSIQ slightly below cutoff were determined eligible by principal investigators using clinical judgment on a case-by-case basis.

on the nine items of the CASS measure, and then to reliability on the confederate adherence code, reflected in at least 80% agreement with trained research personnel across 10% of coded videos.

Conversation Rating Scale (CRS; Ratto et al., 2011)—The CRS, designed to evaluate the confederate's behavior (i.e., engagement) in the conversation, was completed twice by participants; once after CASS-I and again after CASS-B. The CRS was also completed by the study's trained coder, where reliability was determined as at least 80% agreement with trained research personnel across 10% of coded videos. The five items of the CRS are rated on a 7-point Likert scale measuring agreement (e.g., Strongly Disagree… Strongly Agree) and totaled, wherein larger values reflect the participant's perception of the confederate as more engaged in the conversation.

Autism Diagnostic Observation Schedule-2nd edition (ADOS-2, Lord et al., 2012)—The ADOS-2 Module 3 is a gold standard, diagnostic assessment of ASD symptoms which includes structured and unstructured situations that allow for the observation of spontaneous social-communicative behavior. Trained clinicians code these behaviors, where greater scores on the ADOS-2 reflect greater impairment across two subscales of Social Affect (ADOS-SA; i.e., social communication and reciprocal social interaction), and Repetitive Behaviors. The ADOS-SA offers particular utility as a clinician rating of observed social behavior.

Social Responsiveness Scale, Second Edition (SRS-2; Constantino & Gruber, 2012)—The SRS-2 is a parent-report survey designed to measure symptoms of ASD in school-age adolescents with strong internal consistency (α=.95, Constantino & Gruber, 2012). Each of the 65 items is rated on a 4-point Likert scale, (not true – always true); raw scores are totaled and converted to T-scores, where greater scores suggest greater impairment. The present study utilized the Social Communication Index (SRS-SCI), which includes the social cognition, social motivation, social awareness, and social communication subscales as an index of parent-reported social behavior.

NEPSY-II (Korkman, et al., 2007)—The NEPSY-II is a collection of assessments designed to evaluate neuropsychological development in youth aged 3–16. Four assessments were used in the present study as measurements of social cognition. Raw scores on each subtest are standardized by age, and the measure has shown high internal reliability across subtests (Korkman, et al., 2007).

NEPSY Affect Recognition Task (AR) subtest includes 33 items evaluating the ability to identify facial expressions of various emotions. Total raw scores are standardized, where greater raw scores and scaled scores reflect increased accuracy in identifying expressed emotions.

NEPSY Memory for Faces (MF) and *NEPSY Memory for Faces Delayed* (MFD) subtests evaluate the ability to identify sixteen recently-presented images of faces immediately after presentation and after a 15–25 minute delay. Total raw scores are standardized, where greater raw scores and scaled scores reflect increased ability to recall familiar faces.

NEPSY Theory of Mind subtest includes twenty-eight items evaluating the ability to interpret and predict the behavior of others across a variety of scenarios and with varying information. Total raw scores are standardized, where greater raw scores and scaled scores reflect increased ability to apply mental concepts to interpret others' behavior.

Data Analytic Plan

All statistical analyses were conducted using SPSS Statistics Software, version 25. Preliminary analyses of primary study variables were conducted prior to addressing the study's main aims in order to examine data distribution and confirm all assumptions for statistical analyses had been met. A χ^2 test of independence was conducted to determine whether confederates adhered to their prescribed condition, per hypothesis 1. To address hypothesis 2 and examine changes in participant behavior between CASS-I and CASS-B, we conducted a repeated measures ANOVA. Of note, two items (Questions Asked; Topic Changes) are count variables, which should be considered when interpreting this model (Jaeger, 2008)[3]. The CASS Social Adaptation Index (SAI) was also computed, wherein primary outcomes of Questions Asked, Topic Changes, Overall Involvement, and Overall Quality of Rapport in CASS-I are subtracted from CASS-B to create a composite difference score (White et al., 2015). The SAI is thought to reflect normative social adaptation, where individuals will exhibit an increase in Questions Asked, Topic Changes, and Overall Involvement and a decrease in Overall Rapport between CASS-I and CASS-B. To test hypothesis 3, the CRS-C score was computed as a metric of participant insight into confederate behavior, and a paired-samples *t*-test was used to compare participant ratings of interested and bored confederates. Per hypothesis 4, a repeated measures ANOVA was used to compare participant and coder ratings of confederate behavior during CASS-I and CASS-B. Bivariate and partial Pearson's *r* correlations, adjusted for multiple comparisons (Benjamini & Hochberg, 1995), were used to evaluate hypothesis 5 and analyze convergent validity of the CASS and accompanying CRS.

Results

Preliminary Analyses and Descriptive Data

All 50 participants completed the CASS, SRS-2, and NEPSY tasks. Reliability analyses of the CASS measure suggested sufficient internal consistency across all subscales (Table 2), with the exception of the CRS-Interested subscale (α=.56). Closer scrutiny revealed removal of the fifth item ("the person created a sense of distance between us") improved the overall reliability of the scale (α=.71), suggesting participants may not have understood this question. All analyses were subsequently run with a shortened CRS excluding Item 5; unless noted, there were not significant differences in results when this item was omitted. To avoid ceiling effects and better capture variance across measures, all analyses were conducted using raw scores from SRS-2 and NEPSY measures.

[3]General Estimating Equations (GEE) were initially conducted (Proudfoot, Lin, Wang, & Tu, 2018). Results indicated that an independent correlation structures with a Poisson distribution provided a superior fit to other models (e.g., normal distribution; unstructured correlation structure). However, the pattern of results for these models did not differ from the original repeated measures ANOVA; thus, the original model is reported for clarity, simplicity, and replicability.

Aim 1. Examining Confederate Adherence

Overall, 79% of participants adhered to their prescribed condition; however, a χ^2 test of independence reflected a significant difference in these categorical ratings across conditions, where 96% of CASS-I conversations were rated as 'hit' compared to 62% of CASS-B conversations ($\chi^2(100)=17.42$, $p< .001$). Adherence ratings did not differ by site or gender. Findings do not fully support hypothesis 1; <80% of all confederates adhered to their prescribed condition, and bored confederates had particular difficulty with adherence.

Aim 2. Examining Internal and Convergent Validity of the CASS

To examine differences in participants' behavior when talking with interested versus bored confederates, each of the nine CASS-I and CASS-B items were using a repeated measures ANOVA (Table 3). The overall test was significant (Wilk's $\Lambda=.48$, $F(9,41)=4.90$, $p<.001$, partial $\eta^2=.52$), and pairwise comparisons indicated participants asked significantly more questions of bored confederates than interested confederates ($F(1)= 23.26$, $p<.001$, partial $\eta^2=.32$) and overall quality of rapport was significantly better during CASS-I ($F(1)=8.88$, $p=.004$, partial $\eta^2=.15$), providing partial support for hypothesis 2.

A paired samples *t*-test was used to examine differences in CRS ratings of interested and bored conversation partners. Overall, participants rated interested confederates as significantly more engaged than bored confederates ($t(49)= 4.06$, $p<.001$, Cohen's $d=.774$, Figure 2). A change score (CRS-C) was computed by subtracting CASS-B CRS ratings from CASS-I CRS ratings, where positive values reflect participants' identification of the interested confederate as more engaged than the bored confederate (see Leonczyk, 2017). A frequency analysis indicated seven participants (14%) rated bored confederates as more engaged; seven rated the two conversation partners equally (14%), and the remaining 36 participants (72%) indicated the interested confederate was more engaged than the bored. An independent samples *t*-test indicated the average CRS-C score of 4.06 was significantly different from 0 ($t(49)= 5.48$, $p<.001$), suggesting participants were generally able to distinguish between interested and bored conversation partners, supporting hypothesis 3.

Participant and trained coder CRS ratings were compared using a repeated measures ANOVA. The overall test was significant (Wilk's $\Lambda=.48$, $F(2,48)= 25.92$, $p<.001$, partial $\eta^2=.52$), suggesting participants' ratings of interested and bored confederates were more neutral than the trained coder's ratings; CASS-I: $F(1)= 4.89$, $p=.032$, partial $\eta^2=.091$; CASS-B: $F(1)= 46.40$, $p < .001$, partial $\eta^2=.486$ (Figure 2). While overall comparison of ratings omitting item 5 from the CRS was still significant (Wilk's $\Lambda=.45$, $F(2,48)= 29.46$, $p<.001$, partial $\eta^2=.55$), participants' ratings of CASS-I confederates did not significantly differ from the trained coder's ratings ($F(1)=.39$, $p=.537$, partial $\eta^2=.008$). The large effect sizes capturing the difference between ratings of bored confederates suggests participants were particularly unaware of confederates' disinterest, supporting hypothesis 4. Pearson correlations were used to further examine these associations. Despite the setup of the CASS, which intends to provide the participant with two distinct conversational experiences, participants' CRS ratings across CASS-I and CASS-B conversations were significantly related ($r=.47$, $p=.001$), whereas such an association was not found between the trained coder's ratings of the two conditions ($r= -.20$, $p=.160$). This difference, although not

statistically significant (Fisher's r to z transformations $p > .05$), suggests participants may not have perceived differences between the two conversations as strongly as did the coders. Alternatively, perhaps the ASD participants rate conversations more similarly for other reasons. A single, trained coder rated each video, therefore this difference is not attributable to actor variance and instead reflects a relative deficit in participants' ability to discriminate between bored and interested confederates, again supporting hypothesis 4.

Participants' CRS ratings and coded behavior during the CASS were then examined for convergent validity (Table 4). After correcting for multiple comparisons (Benjamini & Hochberg, 1995), SAI association with CASS-B composite was marginally significant (r=.34, p=.067) and association with CASS-I composite was nonsignificant, suggesting that CASS-B accounts for more variance in SAI scores. SAI was not associated with any other measures of social behavior or cognition. CASS-I composite was significantly correlated with CASS-B composite (r=.72, p<.001) and ADOS-SA (r= −.44, p=.02), while CASS-B composite association with ADOS-SA was only marginally significant (r=.36, p=.06). SRS-SCI was not associated with any CASS indices. Overall findings demonstrate convergent validity of CASS-I, but not of CASS-B or SAI, as an assessment of social behavior; participants' behavioral shift (i.e., SAI) did not account for any additional variance in measured social behavior. CRS-C scores were moderately associated with CASS-I composite (r=.44, p=.017), while marginally significant associations were noted with ADOS-SA (r= −.38, p=.057 and all four NEPSY social cognitive tasks (r=.34–.37, p=.051–.064). These findings support hypothesis 5, suggesting participants' rating of confederate behavior is associated with extant measures of social cognition. The trained coder's CRS-C score was included as a covariate in a series of partial correlations to control for nonadherent confederate behavior; comparison of findings suggested this adjustment did not significantly impact results, nor did excluding CASS conversations with nonadherent confederates (Fisher's r to z transformations $p > .05$, data available upon request).

Discussion

We sought to examine confederate behavior during the CASS and validate the CASS itself by anchoring scores to extant measures of social cognition and social behavior. To these ends, CASS behavior codes of participants with ASD and confederates were analyzed along with the CRS and other measures of social cognition and behavior. Findings support for the utility of the CASS as an ecologically sound measure of broader social competence, assessing both social cognition and social behavior in youth with ASD. The CRS appears especially useful as a measure of social cognition for verbally fluent individuals, as it highlights participants' insight into their conversation partner's social behavior and can be used as a referent for changes in their own social behavior. The observed changes in participant behavior across conditions, and noted utility of the CRS-C as an assessment of social cognition, emphasize the importance of the bored condition of the CASS.

Although participants' CRS ratings suggested they were able to discriminate between interested and bored confederates, corroborating work done by Ratto and colleagues (2011), these ratings were significantly related and also differed significantly from a trained coder's CRS ratings, suggesting that participants were relatively less aware of, or

perhaps relatively less confident in, the differences between their conversation partners. These findings are consistent with prior research noting impaired cognition in youth with ASD (Church et al., 2000; Golan et al., 2006). Bored confederates appeared to struggle, in particular, with the instructions for the CASS, which likely reflects difficulty and discomfort with pretending to be disinterested (i.e., defying rules of social etiquette) during a social conversation. However, despite relatively poor adherence in this condition, including confederate adherence as a covariate did not significantly alter the pattern of findings, which suggests the CASS is robust to confederate behavior drift.

The CRS-C shows promise as a metric of social cognition, as it was associated with all four NEPSY assessments of social cognitive ability. Meanwhile, the CASS-I composite score was significantly, inversely related to the ADOS-SA assessment of social behavior, but this convergent validity was not found with the SAI or CASS-B composite. Importantly, the SRS-2 asks for behavior ratings “over the last six months” which may not directly relate to expressed behavior during the clinical assessment, accounting in part for the nonsignificant associations with CASS variables. Prior research using the bored condition has not examined the CASS-B composite as an independent index (e.g., White et al., 2015), likely because talking with a confederate instructed to act disinterested may not generalize to typical social interactions. Findings from this study suggest the CASS-B, does not, in fact, align with established metrics of social behavior. Nevertheless, additional research is needed to clarify how parent report, which is most typically used to assess and evaluate clinical interventions of social competency in ASD (Gates et al., 2017), might fit into this pattern.

Interestingly, the CASS-I composite score of social behavior was modestly correlated with the CRS-C, while the comprehensive adaptation score (SAI) and CASS-B composite score were unrelated to CRS-C. Additional research is needed to argue whether the SAI, which captures behavioral change between the interested and bored CASS conversation, is superior to the CASS-I as tool to measure social behavior. It may be the SAI is an inadequate composite of behavioral change, or it may be the SAI is, instead, the purest metric of social behavior specifically because it is unrelated to social cognition. Some work in the neurotypical population indicates a relationship between social cognition and social behavior (e.g., Ford, 1982), but these constructs are complex and multifaceted. Other work demonstrates nuanced distinctions; Barendse and colleagues (2018) found adolescents with high-functioning ASD performed no differently than controls on measures of social cognition, yet still demonstrated social functioning deficits in everyday life. Likewise, some research indicates knowledge of social rules or skills does not always translate to social behavior, in both the ASD (Hill & Frith, 2003) and typically developing population (Lerner et al., 2012). The current findings emphasize this distinction between social cognition and behavior and highlight the importance of considering both of these constructs in outcome measures and assessment.

Limitations

This study’s sample was relatively homogeneous (86% Caucasian; 100% non-Hispanic); future work should aim to validate these findings in a more diverse population. The use and interpretation of difference scores (e.g., the Social Adaptation Index and CRS-

change score) has been challenged recently, as these scores may not reflect any additional information (i.e. variance) not already conveyed in the examined components (Laird, 2020). These concerns bear consideration in continued efforts to understand and replicate noted relationships between the CASS, CRS, and extant measures of social behavior and cognition. Additionally, this study examined only the ADOS-2, SRS-2 and NEPSY as measures of social behavior and social cognition, respectively. Although these measures are widely used and normed, examining convergent validity using just these measures may not fully capture the extent to which the CASS taps the constructs of social behavior and social cognition. Moreover, this study does not consider social motivation, which may affect the extent to which an individual with ASD acts upon their awareness of a conversation partner (e.g., he/she may not recognize the confederate is bored or fail to modify behavior in the CASS-B simply because he/she is disinterested in the conversation). In light of these findings, future research should explore models integrating measurement tools which paint a more accurate picture of the complex association between social behavior and social cognition in youth with ASD.

Future Directions

The current work supports continued investigation into the use of the CASS as a tool capturing components of both social cognition and social behavior. While the CASS requires notably more resources to implement than a standardized parent survey, findings support the use of both the CASS-I and CASS-B, along with the accompanying CRS, for social intervention research seeking to capture both social cognition and observable social behavior with relative confidence in the ecological validity of their findings. Some recent research (e.g., PEERS® program, Laugeson, Gantman, Kapp, Orenski, & Ellingsen, 2015), has opted to exclude CASS-B (e.g., Rabin et al., 2018) for practical reasons or because skills taught in intervention might actually make the client look ‘worse’ during the interested condition (e.g., if client exits the conversation). Parallel research has also integrated unstructured social conversations as a metric of social behavior rather than the structured CASS (e.g., Ko et al., 2019), allowing for examination of a potentially more authentic conversation between peers. In both cases, however, eliminating engagement with both an interested and bored conversation partner forfeits the opportunity to assess social cognition, which hinges on the implementation of CASS-B and accompanying CRS. Future work should continue to explore the potential impact of repeated administration of CASS-B on participants’ behavior (e.g., at pretreatment and posttreatment visits), paying particular attention to the potential impact on ecological validity if the confederate were to be recognized as part of the assessment. Finally, continued research should examine the extent to which social cognition, social motivation, and cognitive ability (i.e., FSIQ) are interrelated and account for variance in social behavior. In the current study, relationships between NEPSY Social Cognitive Tasks and CRS-C scores were nonsignificant when controlling for participants’ FSIQ, which likely reflects the strong association between NEPSY variables and FSIQ (data available upon request). Though beyond the scope of this investigation, findings from such research would inform intervention work utilizing the CASS in clinical research settings.

Conclusions

Findings suggest the CASS effectively parses social cognition, measured using the Conversation Rating Scale, from observable social behavior, measured using the CASS-Interested composite score and the Social Adaptation Index. Moreover, differences in ratings of confederate behavior suggest the CASS is sensitive to deficits in social cognition in participants with ASD. Participants' perceived differences in behavior between interested and bored confederates correlated with existing measures of social cognition, supporting the use of the CASS as a metric of sociocognitive ability, specifically perception of others' social cues. Our finding that CASS social behavior metrics did not consistently correlate with social cognition metrics highlights distinctions between cognition and behavior. These results also encourage careful training and monitoring of confederates, as acting disinterested seems to be challenging for most confederates. Overall, findings support utility of the CASS as a measure of social behavior *and* social cognition in ASD and lay groundwork for further exploration into these distinct components of social competence.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to acknowledge the research assistants in the Vanderbilt SENSE Lab (Director: Blythe Corbett, Ph.D.), the Stony Brook Social Competence and Treatment Lab (Director: Matthew Lerner, Ph.D.), and at The University of Alabama's Center for Youth Development and Intervention (Director: Susan White, Ph.D.) for their assistance with recruitment and data collection. This study was funded by the National Institute of Mental Health (Grant Number: MH114906, PI: Corbett). This research was prospectively reviewed and approved by the Institutional Review Boards (IRB) at Vanderbilt University, The University of Alabama, and Stony Brook University, managed through IRB Reliance Exchange, a web-based portal supporting single IRB review documentation and coordination for multi-center clinical trials. The authors declare no conflicts of interest.

References

Allen M (2017). The Sage Encyclopedia Of Communication Research Methods (Vols. 1–4). Thousand Oaks, CA: SAGE Publications, Inc doi: 10.4135/9781483381411

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders (5th ed.). Washington, DC: American Psychiatric Association.

Barendse EM, Hendriks MPH, Thoonen G, Aldenkamp AP, & Kessels RPC (2018). Social behaviour and social cognition in high-functioning adolescents with autism spectrum disorder (ASD): Two sides of the same coin? Cognitive Processing, 19(4), 545–555. doi:10.1007/s10339-018-0866-5 [PubMed: 29959562]

Benjamini Y & Hochberg Y (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological), 57(1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

Chevallier C, Kohls G, Troiani V, Brodkin ES, & Schultz RT (2012). The social motivation theory of autism. Trends in Cognitive Sciences, 16(4), 231–239. doi:10.1016/j.tics.2012.02.007 [PubMed: 22425667]

Church C, Alisanski S, & Amanullah S (2000). The Social, behavioral, and academic experiences of children with Asperger syndrome. Focus on Autism and Other Developmental Disabilities, 15(1), 12–20. doi:10.1177/108835760001500102

Constantino JN, & Gruber C (2012). Social Responsiveness Scale (2nd ed.). Western Psychological Services.

Corbett BA, Key AP, Qualls L, Fecteau S, Newsom C, Coke C, & Yoder P (2016). Improvement in social competence using a randomized trial of a theatre intervention for children with autism spectrum disorder. Journal of Autism and Developmental Disorders, 46(2), 658–672. doi:10.1007/s10803-015-2600-9 [PubMed: 26419766]

Cordier R, Speyer R, Chen Y-W, Wilkes-Gillan S, Brown T, Bourke-Taylor H, Doma K, & Leicht A (2015). Evaluating the psychometric quality of social skills measures: A systematic review. PLOS ONE, 10(7), e0132299. doi:10.1371/journal.pone.0132299 [PubMed: 26151362]

Dolan BK, Van Hecke AV, Carson AM, Karst JS, Stevens S, Schohl KA, Potts S, Kahne J, Linneman N, Remmel R, & Hummel E (2016). Brief Report: Assessment of intervention effects on in vivo peer interactions in adolescents with autism spectrum disorder (ASD). Journal of Autism and Developmental Disorders, 46(6), 2251–2259. doi:10.1007/s10803-016-2738-0 [PubMed: 26886470]

Ford ME (1982). Social cognition and social competence in adolescence. Developmental Psychology, 323–340.

Gates JA, Kang E, & Lerner MD (2017). Efficacy of group social skills interventions for youth with autism spectrum disorder: A systematic review and meta-analysis. Clinical Psychology Review, 52, 164–181. doi:10.1016/j.cpr.2017.01.006 [PubMed: 28130983]

Golan O, Baron-Cohen S, & Hill J (2006). The Cambridge Mindreading (CAM) Face-Voice Battery: testing complex emotion recognition in adults with and without Asperger Syndrome. Journal of Autism and Developmental Disorders, 36(2), 169–183. doi:10.1007/s10803-005-0057-y [PubMed: 16477515]

Hill EL, & Frith U (2003). Understanding autism: Insights from mind and brain. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358(1430), 281–289. doi:10.1098/rstb.2002.1209 [PubMed: 12639326]

Jaeger TF (2008). Categorical data analysis: Away from ANOVAs. Journal of Memory and Language, 59(4), 434–446. doi:10.1016/j.jml.2007.11.007. [PubMed: 19884961]

Kennedy DP, & Adolphs R (2012). The social brain in psychiatric and neurological disorders. Trends in Cognitive Sciences, 16(11), 559–572. doi:10.1016/j.tics.2012.09.006 [PubMed: 23047070]

Ko JA, Miller AR, & Vernon TW (2019). Social conversation skill improvements associated with the Social Tools And Rules for Teens program for adolescents with autism spectrum disorder: Results of a randomized controlled trial. Autism, 23(5), 1224–1235. doi:10.1177/1362361318808781 [PubMed: 30378448]

Korkman M, Kirk U, & Kemp S (2007). NEPSY-Second Edition (NEPSY II). Harcourt Assessment.

Laird RD (2020). Analytical challenges of testing hypotheses of agreement and discrepancy: Comment on Campione-Barr, Lindell, and Giron (2020). Developmental Psychology, 56(5), 970–977. doi:10.1037/dev0000763 [PubMed: 32271076]

Laugeson EA, Gantman A, Kapp SK, Orenski K, & Ellingsen R (2015). A randomized controlled trial to improve social skills in young adults with autism spectrum disorder: The UCLA PEERS® Program. Journal of Autism and Developmental Disorders, 45(12), 3978–3989. doi:10.1007/s10803-015-2504-8 [PubMed: 26109247]

Lemerise EA, & Arsenio WF (2000). An integrated model of emotion processes and cognition in social information processing. Child Development, 71(1), 107–118. doi:10.1111/1467-8624.00124 [PubMed: 10836564]

Leonczyk C (2017). The effectiveness and feasibility of the PEERS® school-based social skills intervention for adolescents with social difficulties. The University of Alabama at Birmingham.

Lerner MD, White SW, & McPartland JC (2012). Mechanisms of change in psychosocial interventions for autism spectrum disorders. Dialogues in Clinical Neuroscience, 14(3), 307–318. [PubMed: 23226955]

Lord C, Rutter M, DiLavore PC, Risi S, Gotham K, & Bishop S (2012). Autism Diagnostic Observation Schedule, Second Edition (ADOS-2). Torrance, CA: Western Psychological Services, 2012.

McMahon CM, Lerner MD, & Britton N (2013). Group-based social skills interventions for adolescents with higher-functioning autism spectrum disorder: A review and looking to the future. *Adolescent Health*, Medicine and Therapeutics, 4, 23. doi: 10.2147/AHMT.S25402

Pallathra AA, Calkins ME, Parish-Morris J, Maddox BB, Perez LS, Miller J, Gur RC, Mandell DS, Schultz RT, & Brodkin ES (2018). Defining behavioral components of social functioning in adults with autism spectrum disorder as targets for treatment: Components of social functioning in adult ASD. Autism Research, 11(3), 488–502. doi:10.1002/aur.1910 [PubMed: 29341497]

Proudfoot JA, Lin T, Wang B, & Tu XM (2018). Tests for paired count outcomes. General Psychiatry, 31(1), 2–7. doi:10.1136/gpsych-2018-100004

Rabin SJ, Israel-Yaacov S, Laugeson EA, Mor-Snir I, & Golan O (2018). A randomized controlled trial evaluating the Hebrew adaptation of the PEERS ® intervention: Behavioral and questionnaire-based outcomes: RCT of the Hebrew adaptation of PEERS®. Autism Research, 11(8), 1187–1200. doi:10.1002/aur.1974 [PubMed: 30095232]

Ratto AB, Turner-Brown L, Rupp BM, Mesibov GB, & Penn DL (2011). Development of the Contextual Assessment of Social Skills (CASS): A role play measure of social skill for individuals with high-functioning autism. Journal of Autism and Developmental Disorders, 41(9), 1277–1286. doi:10.1007/s10803-010-1147-z [PubMed: 21287253]

Sparrow SS, Cicchetti DV, & Salunier CA (2016). Vineland Adaptive Behavior Scales, Third Edition (Vineland-3). San Antonia, TX: Pearson.

Velikonja T, Fett A-K, & Velthorst E (2019). Patterns of nonsocial and social cognitive functioning in adults with autism spectrum disorder: A systematic review and meta-analysis. JAMA Psychiatry, 76(2), 135. doi:10.1001/jamapsychiatry.2018.3645 [PubMed: 30601878]

White SW, Scarpa A, Conner CM, Maddox BB, & Bonete S (2015). Evaluating change in social skills in high-functioning adults with autism spectrum disorder using a laboratory-based observational measure. Focus on Autism and Other Developmental Disabilities, 30(1), 3–12. doi:10.1177/1088357614539836

Wolstencroft J, Robinson L, Srinivasan R, Kerry E, Mandy W, & Skuse D (2018). A systematic review of group social skills interventions, and meta-analysis of outcomes, for children with high functioning ASD. Journal of Autism and Developmental Disorders, 48(7), 2293–2307. doi:10.1007/s10803-018-3485-1 [PubMed: 29423608]

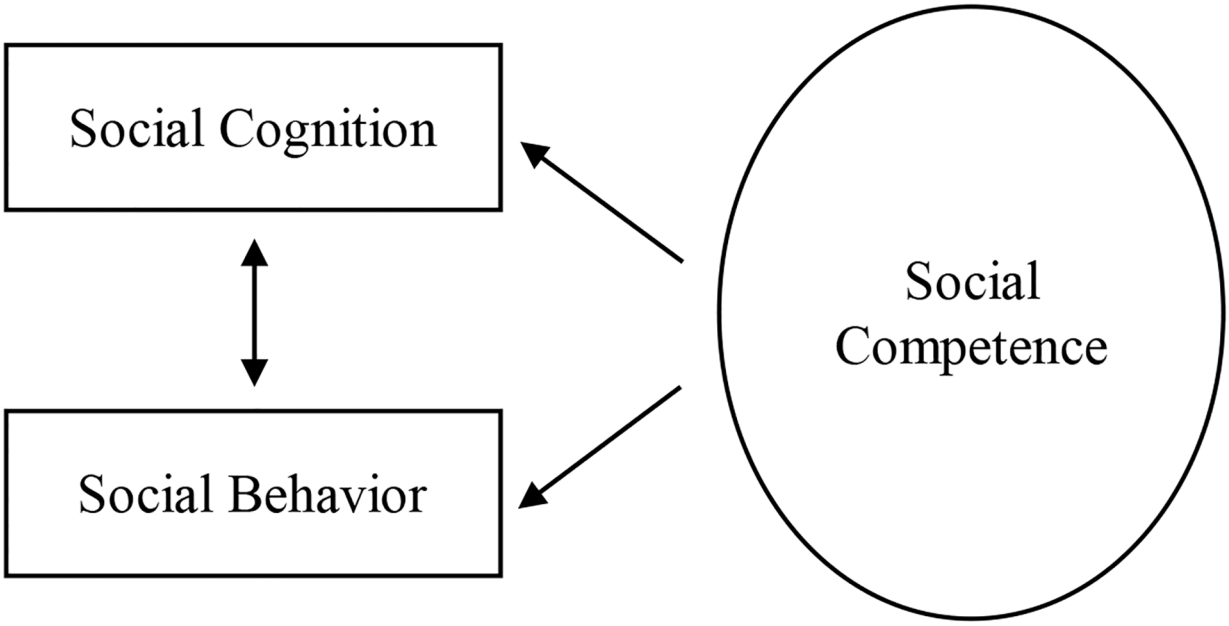

Figure 1.
Proposed Model of Social Competence

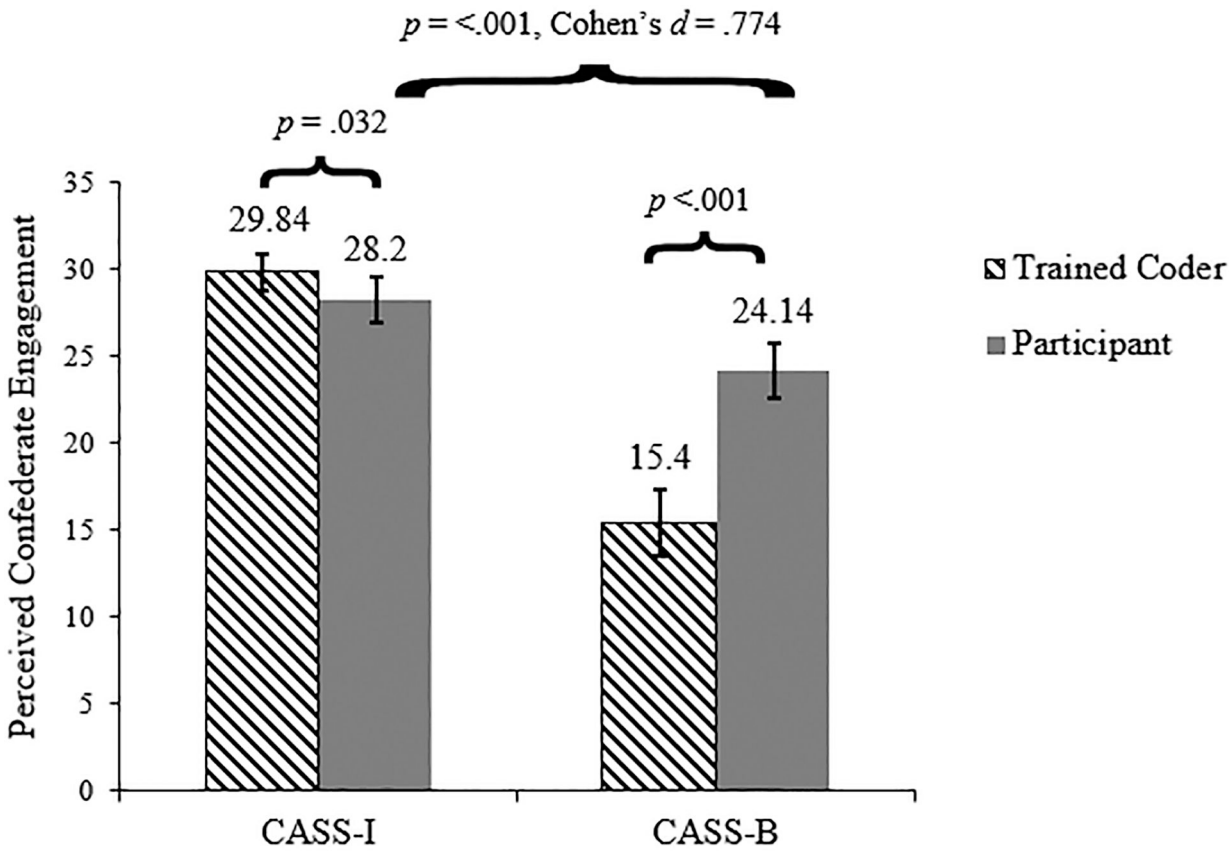

Figure 2.
CRS Ratings of Confederate Behavior

Table 1.

Demographic Data of Participants with ASD

***N* = 50 unless otherwise noted**	**N (%)**		**N (%)**
Race		Diagnosis (*n* = 49)	
White/Caucasian	43 (86%)	ASD	31 (63.3%)
Black/African American	4 (8%)	Autism	12 (24.5%)
American Indian	1 (2)%	Asperger's Syndrome	4 (8.2%)
Asian / Pacific Islander	1 (2)%	PDD-NOS[1]	2 (4.1%)
Other	1 (2%)		
Ethnicity		Gender	
Hispanic	0 (0%)	Male	30 (60%)
Non-Hispanic	50 (100%)	Female	20 (40%)
Parent Education		Family Income	
Less than High School	2 (4%)	Less than $25K	1 (2%)
Graduated from Trade/Business School	1 (2%)	$25K – $50K	13 (26%)
Graduate from High School/GED	7 (14%)	$50K – $75K	5 (10%)
Associates Degree	6 (12%)	$75K – $100K	4 (8%)
Bachelors' Degree	16 (32%)	$100K – $150K	11 (22%)
Master's Degree	9 (18%)	More than $150K	9 (15%)
Doctorate Degree	2 (4%)	Declined / Unknown	7 (14%)
Other	3 (6%)		
Declined to answer	4 (8%)		

	***M* (*SD*)**	**Range**	**Skewness (*SE*)**	**Kurtosis (*SE*)**
Age	12.73 (1.88)	10 – 16	.37 (.34)	−.85 (.66)
FSIQ	98.08 (18.72)	64 – 131	−.07 (.34)	−.85 (.66)

Notes.

[1]Pervasive Developmental Disorder, Not Otherwise Specified

Table 2.

Descriptive Data

N = 50	*M* (*SD*)	Range	Skewness (*SE*)	Kurtosis (*SE*)
CASS Measure				
CASS-I[1] (all 9 items, α = .81)	0.00 (5.88)	−13.67 – 12.24	−.10 (.34)	−.23 (.66)
CASS-B[1] (all 9 items α = .78)	0.00 (5.81)	−14.57— 9.07	−.42 (.34)	−.54 (.66)
CASS-I[1] composite (α = .78)	0.00 (3.27)	−5.74 – 8.62	.35 (.34)	.38 (.66)
CASS-B[1] composite (α = .75)	0.00 (3.27)	−7.36 – 4.91	−.27 (.34)	−1.05 (.66)
SAI[1]	0.00 (1.99)	−3.77 – 5.47	.57 (.34)	.34 (.66)
CRS – Interested (α = .56)	28.20 (4.64)	14 – 35	−.99 (.34)	.97 (.66)
CRS – Bored (α = .72)	24.14 (5.49)	13 – 35	−.03 (.34)	−.78 (.66)
CRS – Interested, Trained Coder (α = .83)	29.84 (3.87)	18 – 35	−.73 (.34)	.66 (.66)
CRS – Bored, Trained Coder (α = .97)	15.40 (7.05)	5 – 34	.59 (.34)	−.24 (.66)
SRS-2 Total Score (T-score)[2]	77.24 (9.51)	53 – 90	−.77 (.34)	.74 (.66)
SRS-SCI (T-score)	76.62 (9.32)	54 – 90	−.54 (.34)	.30 (.67)
ADOS-2 Total Score (n = 49)[3]	12.16 (5.19)	7 – 26	1.20 (.34)	.70 (.66)
ADOS-SA (α = .78)	9.27 (4.24)	4 – 19	.88 (.34)	−.35 (.67)
NEPSY Measure				
Affect Recognition (scaled)	8.42 (3.20)	2 – 16	−.06 (.34)	−.30 (.66)
Memory for Faces (scaled)	7.44 (3.82)	1 – 16	.25 (.34)	−.90 (.66)
Memory for Faces- Delayed (scaled)	7.60 (3.77)	1 – 15	.03 (.34)	−.67 (.66)
Theory of Mind (raw)[4]	21.34 (5.76)	3 – 28	−1.22 (.34)	1.42 (.66)

Notes.

1. Measures converted to Z-scores

2. Table includes scaled scores for clinical interpretation purposes; analyses utilized raw scores

3. ADOS Total score and ADOS-SA subscale score were not available for 1 participant

4. Theory of Mind scaled scores calculated as a percentage; for clarity, raw scores were reported.

CASS: Contextual Assessment of Social Skills; CASS-I: CASS–Interested Composite Score; CASS-B: CASS–Bored Composite Score; SAI: Social Adaptation Index; CRS: Conversation Rating Scale; SRS-SCI: Social Responsiveness Scale–Social Communication Index; ADOS-SA: Autism Diagnostic Observation Schedule-Social Affect Subscale

Table 3.

Repeated Measures ANOVA: CASS-Interested compared to CASS-Bored

	***M(SD)* CASS-I**	***M(SD)* CASS-B**	***F (df = 1)***	***p***	**Partial η^2**
Questions Asked	**3.52 (3.13)**	**5.62 (4.31)**	**23.26**	**<.001**	**.322**
Topic Changes	2.82 (2.28)	3.40 (2.38)	3.35	.073	.064
Involvement	4.34 (1.48)	4.64 (1.51)	2.74	.104	.053
Quality of Rapport	**4.36 (1.21)**	**3.82 (1.27)**	**8.88**	**.004**	**.153**
Vocal Expressiveness	4.82 (1.57)	4.98 (1.49)	1.00	.322	.020
Gestures	3.18 (1.82)	2.96 (1.87)	.87	.356	.017
Positive Affect	4.40 (1.58)	4.32 (1.36)	.22	.642	.004
Kinesic Arousal	3.90 (1.22)	3.92 (1.14)	.02	.900	<.001
Social Anxiety	4.44 (1.43)	4.52 (1.34)	.23	.632	.005

Notes.*M*: Mean; *SD*: Standard Deviation

Table 4.

Bivariate Pearson Correlations (r)

	SAI	CASS-I	CASS-B	CRS-C	SRS-SCI	ADOS-SA (n = 49[a])
CASS-I	−.10	-				
CASS-B	.34[a]	.72***	-			
CRS-C	−.07	.44*	.26	-		
SRS-SCI	−.14	−.09	−.16	−.08	-	
ADOS-SA (n = 49[a])	.10	−.44*	−.36[a]	−.38[a]	−.13	-
Age	−.08	−.03	−.12	.12	.10	.17
FSIQ	.05	.37[a]	.25	.49**	−.06	−.43*
NEPSY-AR	.15	.14	−.02	.35[b]	.03	−.30
NEPSY-MF	−.01	.14	−.05	.37[b]	.04	−.30
NEPSY-MFD	.08	.30	.10	.35[b]	−.01	−.33
NEPSY-TOM	.16	.27	.16	.34[b]	.01	−.30

Notes. Abbreviations: SAI: Social Adaptation Index; CASS-I: CASS–Interested Composite Score; CASS-B: CASS–Bored Composite Score; CRS-C: Conversation Rating Scale–Change Score; SRS-SCI: Social Responsiveness Scale–Social Communication Index; ADOS-SA: Autism Diagnostic Observation Schedule-Social Affect Subscale; FSIQ: Full Scale IQ; NEPSY–AR: Affect Recognition; NEPSY-MF: Memory for Faces; NEPSY-MFD: Memory for Faces Delayed; NEPSY-TOM: Theory of Mind.

[a] ADOS subscale scores were not available for 1 participant, excluded pairwise in correlation analyses

[b] correlations identified as marginally significant, $p < .07$

* $p < .05$,

** $p < .01$,

*** $p < .001$;

Benjamini-Hochberg correction for multiple comparisons